



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/855,294

Source: OIPE

Date Processed by STIC: 5-29-01

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW.

### Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST 25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

## Raw Sequence Listing Error Summary

### ERROR DETECTED

### SUGGESTED CORRECTION

SERIAL NUMBER: 09/855,294

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos     The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will prevent "wrapping."
  
- 2      Invalid Line Length     The rules require that a line **not exceed** 72 characters in length. This includes white spaces.
  
- 3      Misaligned Amino  
    Numbering     The numbering under each 5<sup>th</sup> amino acid is misaligned. Do **not** use tab codes between numbers; use **space characters**, instead.
  
- 4      Non-ASCII     The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. **Please ensure your subsequent submission is saved in ASCII text.**
  
- 5      Variable Length     Sequence(s)      contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the **maximum** number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
  
- 6      PatentIn 2.0  
    "bug"     A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**
  
- 7      Skipped Sequences  
    (OLD RULES)     Sequence(s)      missing. If intentional, please insert the following lines for **each** skipped sequence:  
                               (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                               (i)     SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
                               (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                               This sequence is intentionally skipped  
  
                               Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.
  
- 8      Skipped Sequences  
    (NEW RULES)     Sequence(s)      missing. If **intentional**, please insert the following lines for **each** skipped sequence.  
                               <210> sequence id number  
                               <400> sequence id number  
                               000
  
- 9 ✓      Use of n's or Xaa's  
    (NEW RULES)     Use of n's and/or Xaa's have been detected in the Sequence Listing.  
                               Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
                               In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
  
- 10      Invalid <213>  
    Response     Per 1.823 of Sequence Rules, the only **valid** <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or is Artificial Sequence
  
- 11      Use of <220>     Sequence(s)      missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is **MANDATORY** if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
                               (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
  
- 12      PatentIn 2.0  
    "bug"     Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/855,294

DATE: 05/29/2001

TIME: 11:21:07

Input Set : A:\CRO01-NP007.TXT

Output Set: C:\CRF3\05292001\I855294.raw

4 <110> APPLICANT: Croce, Carlo  
 5 Brenner, Charles  
 6 Fekarski, Yuri  
 10 <120> TITLE OF INVENTION: CRYSTAL STRUCTURE OF WORM NitFhit  
 11 REVEALS THAT A Nit TETRAMER BINDS TWO Phit DIMERS  
 14 <130> FILE REFERENCE: CRO01.NP007  
 C--> 16 <140> CURRENT APPLICATION NUMBER: US/09/855,294  
 C--> 16 <141> CURRENT FILING DATE: 2001-05-15  
 16 <150> PRIOR APPLICATION NUMBER: 60/204,713  
 17 <151> PRIOR FILING DATE: 2000-05-16  
 19 <100> NUMBER OF SEQ ID NOS: 11  
 21 <170> SOFTWARE: FastSEQ for Windows Version 4.0

Does Not Comply  
 Corrected Diskette Needed  
 pp 1-2

## ERRORED SEQUENCES

*Mandatory <220> to <223> features  
 required to explain the "n" at  
 position 1270. see # 9 on the  
 Error Summary Sheet.*

336 <210> SEQ ID NO: 8  
 337 <211> LENGTH: 1359  
 338 <212> TYPE: DNA  
 339 <213> ORGANISM: Homo sapien  
 341 <400> SEQUENCE: 8  
 342 ggggtgcttg tctgcagagt catgacctt ttcgcttgg ccttcattca gcttcagatt 60  
 343 tcttcattca aatcagataa cgtcattcgc gcttctagct ccttcaggga ggcagcaaac 120  
 344 caaggagaca aatcagtttc ttggcaggaa tgccttaatt ccttcattgg agcgaaatat 180  
 345 tttcctgaat atcagagaaa aattccttgg caatccacac agaagcttcc tgaagttaga 240  
 346 aagggaatgca ccttatattt cattggaggg tctatccctg aagaggatgg tgggaaatta 300  
 347 tctaacacct gtctgtgtgt tgggcttcat ggaactttac tagcaaatga tagaaagata 360  
 348 cttctgtttg acattgttgt tcttgcaaaa atttcatttc aagaattctaa aacattgagt 420  
 349 cccggttgata gttctctcac atttgataat ccttactgca gagggtgtgt gggtatctgc 480  
 350 taagacatgc ggttgcaga gcttgcaaaa atttcagcac agagaggttg ccagctgttg 540  
 351 gttatctcag gactttttta ttgacaaat ggaacagccc attgggagtt acttcagaga 600  
 352 agcctgggtg ttgcaatca ggtgtatgtg gcaacagcct ccttcgcccg ggatgacaaa 660  
 353 gctcctatg ttgctggggg acacagcacc gtgggtgaac cttgggggga ggttctagcc 720  
 354 aaagctggca cagcagaagg aatcgtgtat ttagacatag aactgaagaa gctggttgaa 780  
 355 ataagccagg aattccctgt ttttagacag aagcgatcag aactctatgc tctggagatg 840  
 356 aaaaagccct aaagtatttg tttctaatgt gtcacagaa aggaagatat gattctacaa 900  
 357 cataatcaac tctctattaa attctttaat gaagaaaaaa aatttaaaaa aaaaaaaaaa 960  
 358 aactaggttt ccttatggag atgagaaaag cttattatgc tgacatttcc cagccacat 1020  
 359 taaatagtta aaagtatgc agcttgagg cagagagcag aaagctgggc tggttctgaa 1080  
 360 gcttcttcca tacttaagtt gcttcagaag agtttctgag agtatcagat cttggtatcc 1140  
 361 tgggtgattg ttctctaat ataaatatat ttgtgtcatg aactctttaa aaagtctgtg 1200  
 362 gtagttgtaa tctctaat ctaggaaaa gtgggtcttg gtgctattct tttcagaaca 1260  
 E--> 363 ggtacctt(n) aagttccatt tttgggttca tgagtagcta taggaacgca aggggtgatac 1320  
 364 aactttgggt gtttggcag agaagttggg cagccccac 1359  
 396 <210> SEQ ID NO: 13  
 396 <211> LENGTH: 1214  
 397 <212> TYPE: DNA

## RAW SEQUENCE LISTING

DATE: 05/29/2001

PATENT APPLICATION: US/09/855,294

TIME: 11:21:07

Input Set : A:\CRO01-NP007.TXT

Output Set: C:\CRF3\05292001\I855294.raw

398 &lt;213&gt; ORGANISM: X. laevis

400 &lt;400&gt; SEQUENCE: 10

```

401 gtgagtgaacg tttggcgcagt ggcaactaag gcttcctggg aaaatgtaga ggagcgtgag      60
402 cttcgogggga caggacaggg tcttaggcctc tgccttggtg ccacacgccc ttgtgcagac      120
403 tgcctatagac tttgacttta acctgtgtgc cggatatagg ggttagaagc ctgagtgcaa      180
404 tggctgggtgc ccacaagccc ctgattgccc tgtgccagat gacttcaacc ttgataaagg      240
405 agaagaattt cgcacagctgt tggcggtga tccgggaggg tggggggcgt cggcgttga      300
406 tggctgtttct gcggaagcc ttgactata tggggggcag cactgaggag acgctgagtc      360
407 tggctgagtc cctacatggg gacacccatc agcgttacac ccacctggcc agggagtggt      420
408 ggcctctggct ttcctggggg ggatttcctg agaaaggacc caactgggac acggacccac      480
409 gcatttccaa ttctacgtg gttggggaca acacagggca catagtatcg gtgtacggca      540
410 aggcctcaact gtttgacgta gacttgaga atggagtgct actcagagag agcagttcca      600
411 cctcccccgg agcagagctt attcgcccca tcacttctcc agcaggaaaag attggcctgg      660
412 ggggtgtgtta cgaactccgc ttcccagaat tctccttggc tctggcccaa caaggagcag      720
413 aaattctcac ttacccctct gcttcacccc tcaactactg tctggccaat tgggaggtgt      780
414 tgcctgagagc cgttgccata gaaacccagt gctacgtagt tgcagcggca cagacagaca      840
415 gacacaaatga gagaggagc tcttatggc acgctatggt ggtagacccc tgggggctgg      900
416 tcattggcca atgcacaggaa ggaacaggaa tatgttatgc tgagattgac attccctaca      960
417 tggagcgtgt gaggcgggac atgcgcgtgt ggaggcacc caggactgat ctgtatggga      1020
418 aaatctcctt taataaaccc gactgactcc ataatggatc acctgcacct atggggggcaa      1080
E--> 419 acnctttccc ctgattgctg aaattcctca atctgtgact gtgaatgaca atgaacgtga      1140
420 cttggaattg cctgggttat gcacgggcaa tgattctcta cagtaattct caataaaaagt      1200
421 gctgaacctt aaaa      1214

```

<220> to <223> features required  
 to explain the "n" at position  
 1083. # 9 on Error Summary  
 Sheet.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/855,294

DATE: 05/29/2001

TIME: 11:21:08

Input Set : A:\CRO01-NP007.TXT

Output Set: C:\CRF3\05292001\I855294.raw

L:16 M:270 C: Current Application Number differs, Replaced Current Application No

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:363 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:8

L:419 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:10